GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 17, 2005, 18:47:32 ; Search time 137 Seconds
(without alignments)
17.068 Million cell updates/sec

Title: US-10-666-095-6
Perfect score: 23
Sequence: 1 KXVXFXK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1432185 seqs, 334051727 residues

Total number of hits satisfying chosen parameters: 1432185

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	87.0	9	9	US-09-780-053-217	Sequence 217, App
2	20	87.0	9	9	US-09-780-053-326	Sequence 326, App
3	20	87.0	10	9	US-09-780-053-83	Sequence 83, Appl
4	20	87.0	10	9	US-09-780-053-278	Sequence 278, App
5	20	87.0	10	9	US-09-780-053-286	Sequence 286, App
6	20	87.0	10	9	US-09-780-053-300	Sequence 300, App
7	20	87.0	10	9	US-09-780-053-391	Sequence 391, App
8	20	87.0	10	9	US-09-780-053-406	Sequence 406, App
9	20	87.0	10	9	US-09-780-053-709	Sequence 709, App
10	20	87.0	106	11	US-09-864-408A-3306	Sequence 3306, Ap
11	20	87.0	196	16	US-10-767-701-31581	Sequence 31581, A
12	20	87.0	265	14	US-10-032-585-7166	Sequence 7166, Ap
13	20	87.0	328	14	US-10-317-460-10	Sequence 10, Appl
14	20	87.0	328	16	US-10-408-765A-413	Sequence 413, App
15	20	87.0	328	16	US-10-408-765A-1234	Sequence 1234, Ap
16	20	87.0	328	16	US-10-408-765A-2533	Sequence 2533, Ap
17	20	87.0	328	17	US-10-502-279-4	Sequence 4, Appli
18	20	87.0	347	9	US-09-780-053-4	Sequence 4, Appli
19	20	87.0	516	15	US-10-369-493-21928	Sequence 21928, A
20	20	87.0	516	16	US-10-477-369-61	Sequence 61, Appl
21	20	87.0	517	15	US-10-424-599-246165	Sequence 246165,
22	20	87.0	520	15	US-10-425-114-37296	Sequence 37296, A
23	20	87.0	525	15	US-10-425-114-49000	Sequence 49000, A
24	20	87.0	730	9	US-09-780-053-2	Sequence 2, Appli
25	20	87.0	730	14	US-10-145-396-12	Sequence 12, Appl
26	20	87.0	730	14	US-10-409-511-2	Sequence 2, Appli
27	20	87.0	730	17	US-10-726-160-2	Sequence 2, Appli
28	20	87.0	761	16	US-10-437-963-122528	Sequence 122528,
29	19	82.6	44	15	US-10-424-599-194594	Sequence 194594,
30	19	82.6	61	16	US-10-437-963-141529	Sequence 141529,
31	19	82.6	62	15	US-10-424-599-264607	Sequence 264607,
32	19	82.6	69	15	US-10-424-599-211525	Sequence 211525,
33	19	82.6	69	15	US-10-335-977-6346	Sequence 6346, Ap
34	19	82.6	75	15	US-10-424-599-198554	Sequence 198554,
35	19	82.6	105	15	US-10-424-599-143728	Sequence 143728,
36	19	82.6	113	9	US-09-916-790-11	Sequence 11, Appl
37	19	82.6	113	15	US-10-678-786-11	Sequence 11, Appl
38	19	82.6	152	15	US-10-424-599-146591	Sequence 146591,
39	19	82.6	169	15	US-10-424-599-151163	Sequence 151163,
40	19	82.6	184	16	US-10-767-701-31667	Sequence 31667, A
41	19	82.6	186	15	US-10-424-599-210272	Sequence 210272,
42	19	82.6	192	10	US-09-974-879-193	Sequence 193, App
43	19	82.6	192	15	US-10-621-401-193	Sequence 193, App
44	19	82.6	193	10	US-09-305-736-193	Sequence 193, App
45	19	82.6	193	10	US-09-818-683-193	Sequence 193, App

ALIGNMENTS

RESULT 1
US-09-780-053-217
; Sequence 217, Application US/09780053
; Patent No. US20020102640A1

; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 217
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-217

Query Match 87.0%; Score 20; DB 9; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.3e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KXVXFXK 7
| | | |
Db 3 KSVAFSK 9

RESULT 10
US-09-864-408A-3306
; Sequence 3306, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3306
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-3306

Query Match 87.0%; Score 20; DB 11; Length 106;
Best Local Similarity 57.1%; Pred. No. 3e+02;

```
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 KXVXFXK 7
              | | | |
Db         54 KSVAFTK 60


Search completed: May 17, 2005, 19:00:51
Job time : 142 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 17, 2005, 18:40:35 ; Search time 43 Seconds
(without alignments)
12.152 Million cell updates/sec

Title: US-10-666-095-6
Perfect score: 23
Sequence: 1 KXVXFXK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	87.0	180	4	US-09-270-767-57796	Sequence 57796, A
2	20	87.0	223	4	US-09-248-796A-22961	Sequence 22961, A
3	20	87.0	264	4	US-09-248-796A-18523	Sequence 18523, A
4	20	87.0	291	4	US-09-270-767-42499	Sequence 42499, A
5	20	87.0	328	3	US-09-002-298-10	Sequence 10, Appl
6	20	87.0	328	4	US-09-481-277-10	Sequence 10, Appl
7	20	87.0	516	4	US-09-538-092-296	Sequence 296, App
8	19	82.6	81	4	US-09-621-976-7145	Sequence 7145, Ap
9	19	82.6	95	4	US-09-248-796A-21896	Sequence 21896, A
10	19	82.6	97	4	US-09-328-352-7344	Sequence 7344, Ap
11	19	82.6	160	4	US-09-252-991A-24737	Sequence 24737, A

12	19	82.6	184	4	US-09-248-796A-24366	Sequence 24366, A
13	19	82.6	193	3	US-09-041-889-5	Sequence 5, Appli
14	19	82.6	193	3	US-08-837-058-5	Sequence 5, Appli
15	19	82.6	193	4	US-09-417-264-5	Sequence 5, Appli
16	19	82.6	206	4	US-09-270-767-58227	Sequence 58227, A
17	19	82.6	221	3	US-09-247-373B-54	Sequence 54, Appl
18	19	82.6	241	4	US-09-107-532A-4086	Sequence 4086, Ap
19	19	82.6	259	4	US-09-107-532A-5472	Sequence 5472, Ap
20	19	82.6	268	4	US-09-107-532A-5543	Sequence 5543, Ap
21	19	82.6	278	4	US-09-248-796A-20001	Sequence 20001, A
22	19	82.6	282	4	US-09-270-767-42902	Sequence 42902, A
23	19	82.6	292	4	US-09-328-352-4894	Sequence 4894, Ap
24	19	82.6	326	4	US-09-248-796A-15838	Sequence 15838, A
25	19	82.6	344	4	US-09-134-000C-6328	Sequence 6328, Ap
26	19	82.6	358	4	US-09-248-796A-17253	Sequence 17253, A
27	19	82.6	368	4	US-09-248-796A-19795	Sequence 19795, A
28	19	82.6	372	2	US-08-501-003A-12	Sequence 12, Appl
29	19	82.6	383	2	US-08-501-003A-14	Sequence 14, Appl
30	19	82.6	389	2	US-08-501-003A-11	Sequence 11, Appl
31	19	82.6	391	1	US-07-921-178A-2	Sequence 2, Appli
32	19	82.6	391	1	US-08-103-445-5	Sequence 5, Appli
33	19	82.6	391	1	US-08-461-690B-5	Sequence 5, Appli
34	19	82.6	391	2	US-08-501-003A-13	Sequence 13, Appl
35	19	82.6	391	2	US-08-501-003A-16	Sequence 16, Appl
36	19	82.6	391	4	US-09-275-252A-13	Sequence 13, Appl
37	19	82.6	391	4	US-09-949-016-5904	Sequence 5904, Ap
38	19	82.6	398	2	US-08-501-003A-15	Sequence 15, Appl
39	19	82.6	401	4	US-09-248-796A-25547	Sequence 25547, A
40	19	82.6	403	2	US-08-592-383-4	Sequence 4, Appli
41	19	82.6	411	4	US-09-949-016-8100	Sequence 8100, Ap
42	19	82.6	416	3	US-08-764-870-4	Sequence 4, Appli
43	19	82.6	416	3	US-08-980-115-4	Sequence 4, Appli
44	19	82.6	417	4	US-09-134-000C-5002	Sequence 5002, Ap
45	19	82.6	448	4	US-09-949-016-8178	Sequence 8178, Ap

ALIGNMENTS

RESULT 1
US-09-270-767-57796
; Sequence 57796, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57796
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-57796

Query Match 87.0%; Score 20; DB 4; Length 180;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 KXVXFXK 7
              | | | |
Db          4 KSVTFAK 10
```

Search completed: May 17, 2005, 18:48:51
Job time : 45 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 17, 2005, 18:22:23 ; Search time 175 Seconds
(without alignments)
20.483 Million cell updates/sec

Title: US-10-666-095-6
Perfect score: 23
Sequence: 1 KXVXFXK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	87.0	31	2	Q7VPL2	Q7vpl2 haemophilus
2	20	87.0	155	2	Q64MX3	Q64mx3 bacteroides
3	20	87.0	197	2	Q6ZTI9	Q6zti9 homo sapien
4	20	87.0	201	2	Q9C963	Q9c963 arabidopsis
5	20	87.0	211	2	Q96SN0	Q96sn0 homo sapien
6	20	87.0	230	2	Q8RFB3	Q8rfb3 fusobacteri
7	20	87.0	233	2	Q7P7U5	Q7p7u5 fusobacteri
8	20	87.0	283	2	Q9YFL8	Q9yfl8 aeropyrum p
9	20	87.0	288	1	YDIB_SALTY	Q8zpr4 salmonella
10	20	87.0	289	2	O05379	O05379 actinobacil
11	20	87.0	290	2	O66258	O66258 actinobacil
12	20	87.0	300	2	Q8HIU0	Q8hiu0 monosiga br
13	20	87.0	316	2	Q9KL77	Q9kl77 vibrio chol
14	20	87.0	328	1	ECH1_HUMAN	Q13011 homo sapien
15	20	87.0	328	2	Q8WVX0	Q8wvx0 homo sapien

16	20	87.0	328	2	Q96EZ9	Q96ez9 homo sapien
17	20	87.0	376	2	Q6TVD9	Q6tvd9 bovine papu
18	20	87.0	434	2	Q97LQ4	Q97lq4 clostridium
19	20	87.0	469	2	Q9T1T0	Q9t1t0 bacteriopha
20	20	87.0	472	1	GATB_WOLSU	Q7m7y2 wolinella s
21	20	87.0	516	1	TAF6_YEAST	P53040 saccharomyc
22	20	87.0	539	2	Q648V3	Q648v3 uncultured
23	20	87.0	539	2	Q64BE4	Q64be4 uncultured
24	20	87.0	539	2	Q6FYF7	Q6fyf7 bartonella
25	20	87.0	541	2	Q6MDW1	Q6mdw1 parachlamyd
26	20	87.0	558	2	Q9W462	Q9w462 drosophila
27	20	87.0	674	2	Q84N49	Q84n49 zea mays (m
28	20	87.0	711	2	Q6GPU3	Q6gpu3 xenopus lae
29	20	87.0	713	2	Q6P1W0	Q6p1w0 xenopus tro
30	20	87.0	730	2	Q9NWM5	Q9nwm5 homo sapien
31	20	87.0	730	2	Q9NZJ0	Q9nzj0 homo sapien
32	20	87.0	761	2	Q8LH33	Q8lh33 oryza sativ
33	20	87.0	769	2	Q9PLL5	Q9pll5 chlamydia m
34	20	87.0	792	2	Q9M1S4	Q9m1s4 arabidopsis
35	20	87.0	916	2	Q68WR4	Q68wr4 rickettsia
36	20	87.0	1021	2	Q7SA71	Q7sa71 neurospora
37	20	87.0	2011	2	Q8I1Q1	Q8i1q1 plasmodium
38	20	87.0	3381	2	Q8I2V4	Q8i2v4 plasmodium
39	19	82.6	45	2	Q81BJ9	Q81bj9 bacillus ce
40	19	82.6	72	2	Q98485	Q98485 paramecium
41	19	82.6	88	1	SCAB_CANFA	Q95165 canis famil
42	19	82.6	89	2	Q7MR99	Q7mr99 wolinella s
43	19	82.6	91	2	Q74930	Q74930 human immun
44	19	82.6	92	2	Q75852	Q75852 human immun
45	19	82.6	102	2	Q6KI56	Q6ki56 mycoplasma

ALIGNMENTS

```
RESULT 1
Q7VPL2
ID   Q7VPL2      PRELIMINARY;      PRT;    31 AA.
AC   Q7VPL2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=HD0052;
OS   Haemophilus ducreyi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=730;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=35000HP / ATCC 700724;
RA   Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA   Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT   "The complete genome sequence of Haemophilus ducreyi.";
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE017151; AAP95067.1; -.
KW   Complete proteome; Hypothetical protein.
```

SQ SEQUENCE 31 AA; 3634 MW; C4C8055B38CDBFA3 CRC64;

Query Match 87.0%; Score 20; DB 2; Length 31;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 KXVXFXK 7
              | | | |
Db         12 KAVTFTK 18
```

Search completed: May 17, 2005, 18:44:24
Job time : 179 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 17, 2005, 18:32:06 ; Search time 39 Seconds
(without alignments)
17.270 Million cell updates/sec

Title: US-10-666-095-6
Perfect score: 23
Sequence: 1 KXVXFXK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	87.0	201	2	C96634	hypothetical prote
2	20	87.0	283	2	E72780	probable nucleotid
3	20	87.0	290	2	T00111	glycosyltransferas
4	20	87.0	316	2	G82407	D-alanyl-D-alanine
5	20	87.0	328	2	I38882	peroxisomal enoyl-
6	20	87.0	434	2	E96961	membrane protein c
7	20	87.0	516	2	S64120	TATA box-binding p
8	20	87.0	769	2	F81742	conserved hypothet
9	20	87.0	792	2	T47635	probable protein -
10	19	82.6	72	2	T17937	hypothetical prote
11	19	82.6	110	2	D81250	hypothetical prote
12	19	82.6	125	2	T01819	hypothetical prote
13	19	82.6	136	2	E90050	hypothetical prote

14	19	82.6	162	2	G70079	hypothetical prote
15	19	82.6	174	2	H86226	hypothetical prote
16	19	82.6	180	2	T34745	probable proteinas
17	19	82.6	181	2	H81150	hypothetical prote
18	19	82.6	194	1	HSHU10	histone H1-0 - hum
19	19	82.6	208	2	D64380	conserved hypothet
20	19	82.6	233	2	A81945	probable adenosylh
21	19	82.6	271	2	AH1663	amino acid ABC tra
22	19	82.6	283	2	T20367	hypothetical prote
23	19	82.6	298	2	A64058	dihydrodipicolinat
24	19	82.6	306	1	H64539	conserved hypothet
25	19	82.6	315	1	S73917	thioredoxin-disulf
26	19	82.6	326	2	S28706	hypothetical prote
27	19	82.6	327	2	A57626	peroxisomal enoyl
28	19	82.6	333	1	KHRTH	cathepsin H (EC 3.
29	19	82.6	345	2	F97315	uncharacterized co
30	19	82.6	348	2	D70195	hypothetical prote
31	19	82.6	354	2	G84616	hypothetical prote
32	19	82.6	378	2	T34488	hypothetical prote
33	19	82.6	388	2	F70133	flagellar-associat
34	19	82.6	391	2	A55119	potassium channel
35	19	82.6	391	2	S30046	potassium channel
36	19	82.6	406	2	E81300	probable glucose-6
37	19	82.6	409	2	T19326	hypothetical prote
38	19	82.6	420	2	AG1385	B. subtilis YvlB p
39	19	82.6	426	2	D71982	citrate synthase -
40	19	82.6	426	2	B64523	citrate synthase -
41	19	82.6	440	2	JL0144	interleukin-6 rece
42	19	82.6	441	2	E72242	Mg-protoporphyrin
43	19	82.6	443	2	E88343	protein Y38F1A.6 [
44	19	82.6	444	2	I51256	retinoic acid rece
45	19	82.6	448	2	B56558	retinoic acid rece

ALIGNMENTS

RESULT 1
C96634
hypothetical protein T7P1.3 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C96634
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C96634
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <STO>
A;Cross-references: UNIPROT:Q9C963; GB:AE005173; NID:g6751680; PIDN:AAF27663.1; GSPDB:GN00141
C;Genetics:
A;Gene: T7P1.3
A;Map position: 1

Query Match 87.0%; Score 20; DB 2; Length 201;
Best Local Similarity 57.1%; Pred. No. 96;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 KXVXFXK 7
              | | | |
Db         19 KTVAFTK 25
```

Search completed: May 17, 2005, 18:48:03
Job time : 43 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 17, 2005, 18:31:05 ; Search time 164 Seconds
(without alignments)
16.508 Million cell updates/sec

Title: US-10-666-095-6
Perfect score: 23
Sequence: 1 KXVXFXK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	87.0	9	4	AAM24833	Aam24833 Human MHC
2	20	87.0	9	4	AAM24724	Aam24724 Human MHC
3	20	87.0	10	4	AAM24785	Aam24785 Human MHC
4	20	87.0	10	4	AAM24590	Aam24590 Human MHC
5	20	87.0	10	4	AAM25216	Aam25216 Human MHC
6	20	87.0	10	4	AAM24807	Aam24807 Human MHC
7	20	87.0	10	4	AAM24793	Aam24793 Human MHC
8	20	87.0	10	4	AAM24913	Aam24913 Human MHC
9	20	87.0	10	4	AAM24898	Aam24898 Human MHC

10	20	87.0	106	5	ABP32680	Abp32680 Human ORF
11	20	87.0	197	8	ADR08471	Adr08471 Human pro
12	20	87.0	211	4	AAB95316	Aab95316 Human pro
13	20	87.0	265	5	ABP73329	Abp73329 Candida a
14	20	87.0	287	4	ADM19846	Adm19846 Protein e
15	20	87.0	304	8	ADO18819	Ado18819 Human lip
16	20	87.0	328	6	AAE33213	Aae33213 Human mit
17	20	87.0	328	7	ADB80259	Adb80259 PPARgamma
18	20	87.0	328	7	ADC39100	Adc39100 Novel hum
19	20	87.0	328	7	ADE62099	Ade62099 Human Pro
20	20	87.0	328	7	ADH88966	Adh88966 Human per
21	20	87.0	328	7	ADI62982	Adi62982 Human apo
22	20	87.0	328	7	ADJ69428	Adj69428 Human hea
23	20	87.0	328	7	ADJ68607	Adj68607 Human hea
24	20	87.0	328	7	ADJ70727	Adj70727 Human hea
25	20	87.0	328	8	ADF12117	Adf12117 Human per
26	20	87.0	328	8	ADQ30537	Adq30537 Pancreas
27	20	87.0	353	4	ABG08521	Abg08521 Novel hum
28	20	87.0	516	6	ABR52592	Abr52592 Protein s
29	20	87.0	516	7	ADK62574	Adk62574 Disease t
30	20	87.0	516	8	ADS43498	Ads43498 Bacterial
31	20	87.0	558	4	ABB59577	Abb59577 Drosophil
32	20	87.0	673	8	ADQ97697	Adq97697 Human can
33	20	87.0	674	6	ABR41860	Abr41860 Maize gro
34	20	87.0	730	4	AAM25224	Aam25224 Human pro
35	20	87.0	730	6	ABU09611	Abu09611 Human ret
36	20	87.0	730	7	ADF69740	Adf69740 Human ret
37	20	87.0	730	8	ADO20069	Ado20069 Human PRO
38	20	87.0	730	8	ADO20232	Ado20232 Human PRO
39	19	82.6	9	8	ADK65075	Adk65075 PP1c-inte
40	19	82.6	31	6	ABP80894	Abp80894 N. gonorr
41	19	82.6	31	6	ABP77801	Abp77801 N. gonorr
42	19	82.6	31	6	ABP79531	Abp79531 N. gonorr
43	19	82.6	47	3	AAG47978	Aag47978 Arabidops
44	19	82.6	49	4	AAM88805	Aam88805 Human imm
45	19	82.6	67	4	AAO13406	Aao13406 Human pol

ALIGNMENTS

RESULT 1
AAM24833
ID AAM24833 standard; peptide; 9 AA.
XX
AC AAM24833;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human MHC molecule HLA-A11 binding 83P5G4 peptide #10.
XX
KW 83P5G4-related protein; prostate; testis; tissue; cancer; bladder; liver;
KW tumour; kidney; brain; bone; ovary; breast; pancreas; colon; lung; serum;
KW cytostatic; gene therapy; antibody therapy; ribozyme; blood; cervix;
KW single chain monoclonal antibody; urine; uterus; rectum; stomach; human;
KW chromosome 1q31-q32.
XX

OS Homo sapiens.
XX
PN WO200159115-A2.
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US004426.
XX
PR 09-FEB-2000; 2000US-0181261P.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Hubert RS, Afar DEH, Challita-Eid PM, Faris M, Levin E;
PI Mitchell SC, Jakobovits A;
XX
DR WPI; 2001-514669/56.
XX
PT An isolated 83P5G4-related protein useful as a diagnostic and/or
PT therapeutic agent in multiple cancers such as prostate, bladder and bone
PT cancer.
XX
PS Example 15; Page 82; 112pp; English.
XX
CC The polypeptide sequences represent the 83P5G4-related protein and
CC peptide fragments of the protein. 83P5G4 exhibits prostate specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, testis, bladder, kidney,
CC brain, bone, cervix, uterus, ovary, breast, pancreas, stomach, rectum,
CC liver, colon and lung. The 83P5G4 polynucleotide, its related protein and
CC peptide fragments and specific PCR primers are therefore useful for
CC diagnosing and treating cancer. A vector comprising a polynucleotide
CC which encodes a single chain monoclonal antibody, that immunospecifically
CC binds to an 83P5G4-related protein, and a ribozyme capable of cleaving a
CC polynucleotide having the 83P5G4 coding sequence, are both useful in the
CC preparation of a composition for treating a patient with a cancer that
CC expresses 83P5G4. The sequences can be used in diagnostic methods to
CC monitor the level of 83P5G4 gene products in serum, blood, urine and
CC tissue and to thereby detect the presence of cancerous cells
XX
SQ Sequence 9 AA;

Query Match 87.0%; Score 20; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 KXVXFXK 7
              | | | |
Db          3 KSVAFSK 9
```

RESULT 10
ABP32680
ID ABP32680 standard; protein; 106 AA.
XX

AC ABP32680;
XX
DT 08-JUL-2002 (first entry)
XX
DE Human ORF1653 protein, SEQ ID NO:3306.
XX
KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnerary;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
PN WO200190366-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US017076.
XX
PR 24-MAY-2000; 2000US-0206690P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Leach MD, Shimkets RA;
XX
DR WPI; 2002-106200/14.
DR N-PSDB; ABN76706.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX
PS Claim 10; Page 1085; 2508pp; English.
XX
CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,

CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX
SQ Sequence 106 AA;

Query Match 87.0%; Score 20; DB 5; Length 106;
Best Local Similarity 57.1%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 KXVXFXK 7
              | | | |
Db         54 KSVAFTK 60
```

Search completed: May 17, 2005, 18:47:19
Job time : 170 secs